Model Validation & Evaluation: With Whom Lies Responsibility?

Dr. Tom Philp
Oasis Insight
May 2023
New Venture: Maximum Vacuums (MaxVacs)

- MaxVacs offers a new vacuum cleaner to the household market – it’s a great day for everyone.
- Various statements are now being made:
  - ??? has confirmed that the vacuum creates continuous suction
  - ??? has confirmed the suction is sufficiently strong to serve the primary aim of MaxVacs’ target market.
  - ??? has confirmed that the suction is sufficient to pick up pet hair from their £200k Afghan rug
  - ??? has quantified that the vacuum collects 20% more pet hair than the current leading vacuum on the market
  - ??? has quantified that the vacuum performs 50% worse at picking up everyday household dust
  - ??? has proven that the vacuum picks up <99% of the elephant dung at the local zoo
Definitions & Responsibilities

• We desperately need rigour around **definitions** and **responsibilities** in the cat model market – vendors, regulators and users can all help.

• My personal feelings:
  • **VERIFICATION** should be done by **vendors**.
  • “GLOBAL” **VALIDATION** should be done by **vendors and regulators** – but only for an **idealised/average user** in the target market.
  • “LOCAL” **VALIDATION** should be done by **all users (with regulator confirmation)** for their specific use cases.
  • **VALIDATION** criteria should be set by **regulators and independent, risk literate experts**.
  • **VALIDATION** by **vendors** should be produced in a way that allows **user-based inter-comparative evaluations** to be carried out (note – this is almost always impossible at present).
  • **EVALUATION** done by vendors is **always** likely to have a marketing slant.
  • **EVALUATION** should have **measurable, worth-based** output, **VALIDATION** should be **pass vs fail output**.
  • **KEY TAKEAWAY**: Dickie, Matt, Richard and Steve all moonlight as vacuum cleaner testers.